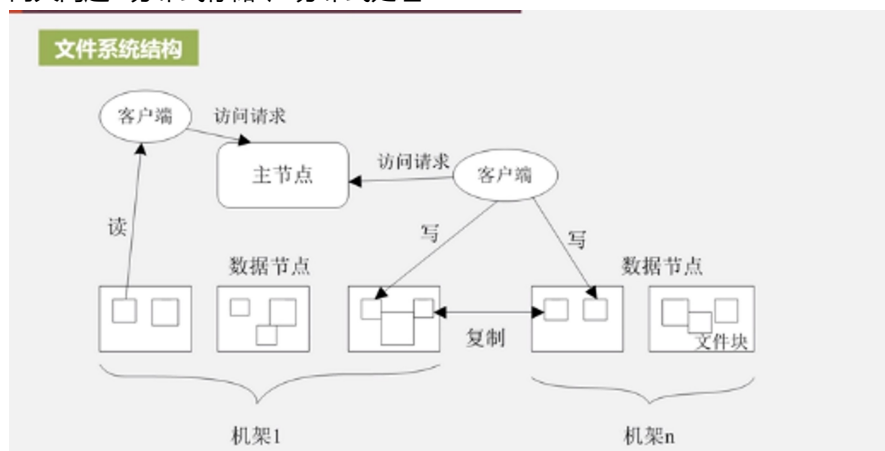


第3讲 分布式文件系统HDFS

2020年3月4日 12:27

全称 Hadoop Distributed File System 简称 HDFS

两大问题：分布式存储、分布式处理



有一个作为主节点（数据目录），从节点

实现目标：兼容廉价的硬件设备、实现流数据读写、支持大数据集（只允许追加不允许修改）、支持简单的文件模型、强大的跨平台兼容性

HDFS自身的局限性

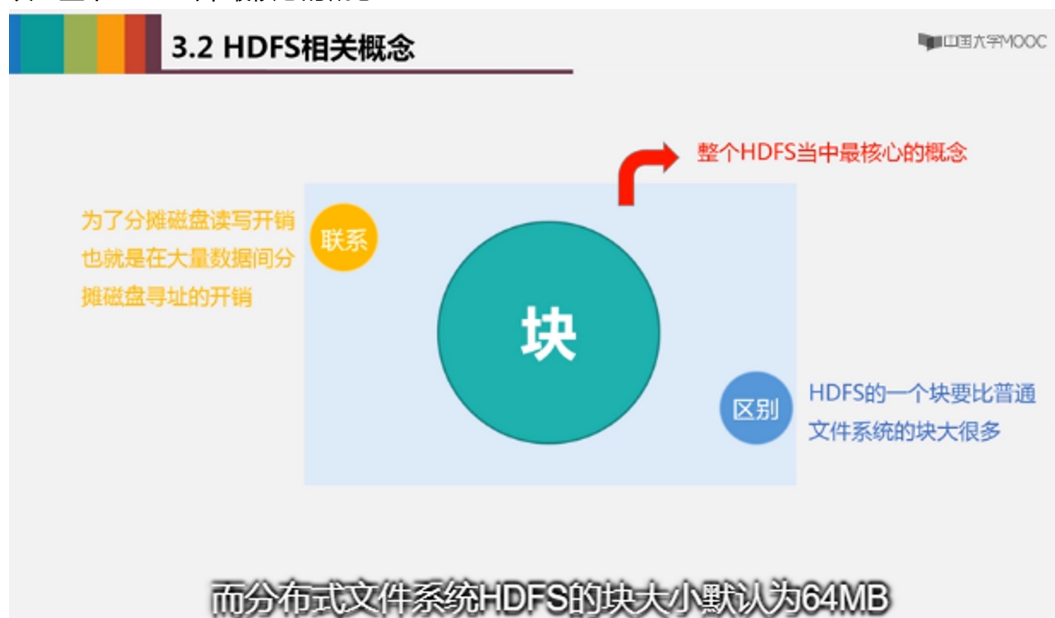
01不适合低延迟数据访问

02无法高效储存大量小文件

03不支持多用户写入及任意修改文件（只允许追加不允许修改）

3.2HDFS相关概念

块：整个HDFS当中最核心的概念



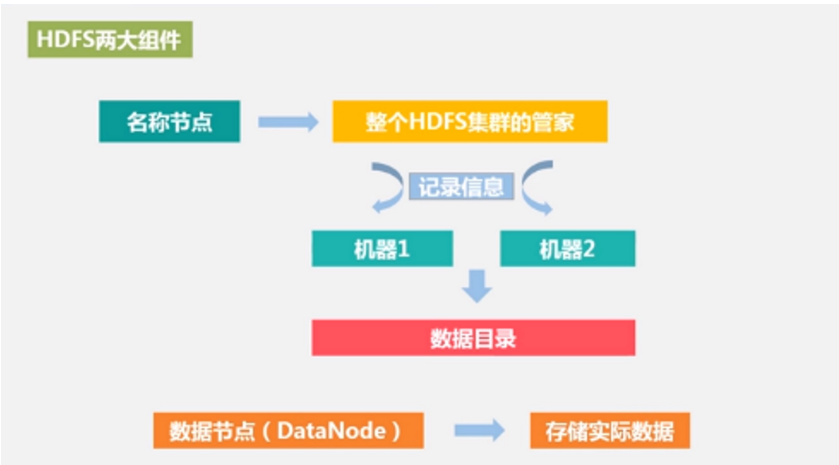
也可设计成128mb

为什么要这么设计? : 1.支持面向大规模数据存储 2.降低分布式节点的寻址开销

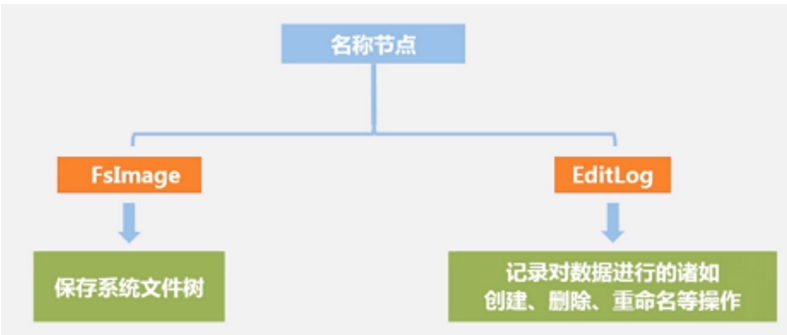
缺点: 如果块过大会导致MapReduce就一两个任务在执行完全牺牲了MapReduce的并行度, 发挥不了分布式并行处理的效果

HDFS采用这种抽象的块的概念设计好处

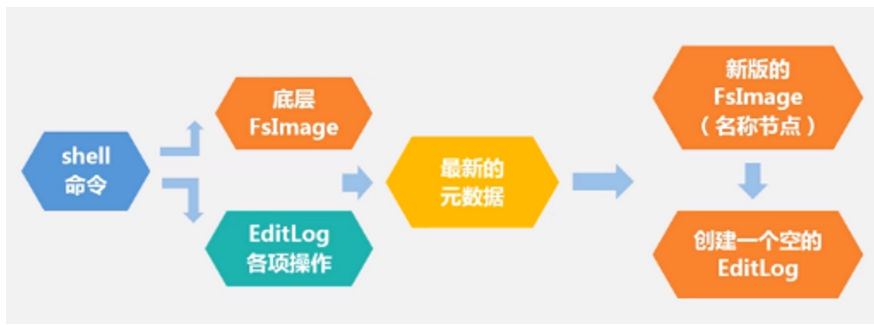
- 1.支持大规模文件存储
- 2.简化系统设计
- 3.适合数据备份



: 数据节点的数据最终都要存储到本地的Linux的文件系统中



FsImage: 记录 文件的复制等级、修改和访问时间、访问权限、块大小以及组成文件的块
数据节点汇报给名称节点 (不断沟通)



Q:为什么有了FsImage还要用EditLog?

A:FsImage对一个HDFS大规模文件系统来讲，文件规模很大，若每次发生数据的更新修改都不断去改FsImage，会导致系统运行非常慢。因此做了一个优化处理，既把更新的部分单独放到EditLog中，不去修改FsImage，把增量后面的更新放到EditLog中，但是EditLog会不断增大（因为会一直有修改动作）。所以有第二名称节点。

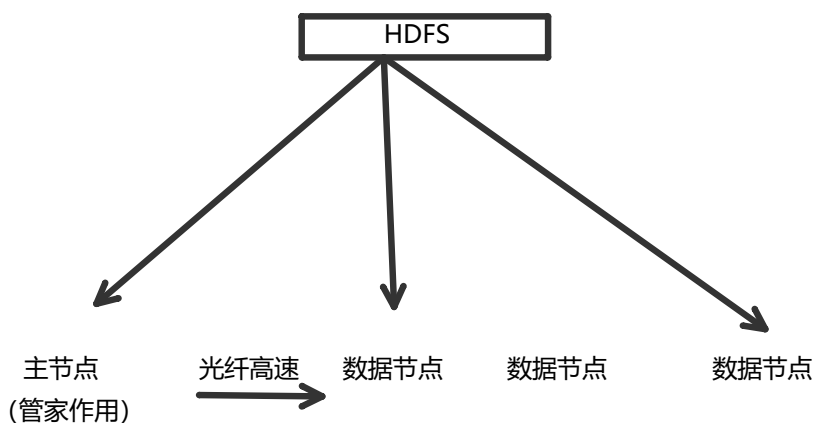
第二名称节点：

- 1.名称节点的冷备份
- 2.对EditLog的处理



数据节点：存、取保存到Linux文件系统中去

3.3HDFS体系结构



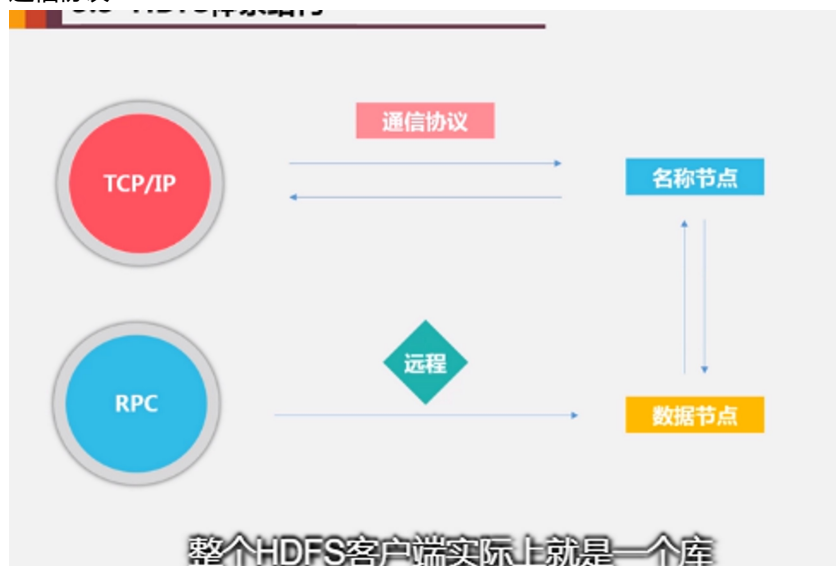
HDFS命名空间

目录 文件 块： /+ 目录名称

访问HDFS文件系统

/usr/shares

通信协议：



不同组件之间通讯有差别。一般可以通过客户端实现打开读取等操作。

HDFS的局限性：

- 1.命名空间限制：名称节点是保存在内存中的，因此，名称节点能够容纳的对象（文件、块）的个数会受到空间大小限制
- 2.性能的瓶颈：整个分布式文件的吞吐量，要受限限于单个名称节点的吞吐量
- 3.隔离问题：由于集群中只有一个名称节点，只有一个命名空间，因此无法对应用程序进行隔离
- 4.集群的可用性：一旦这个唯一的名称节点发生故障，会导致整个集群变得不可用

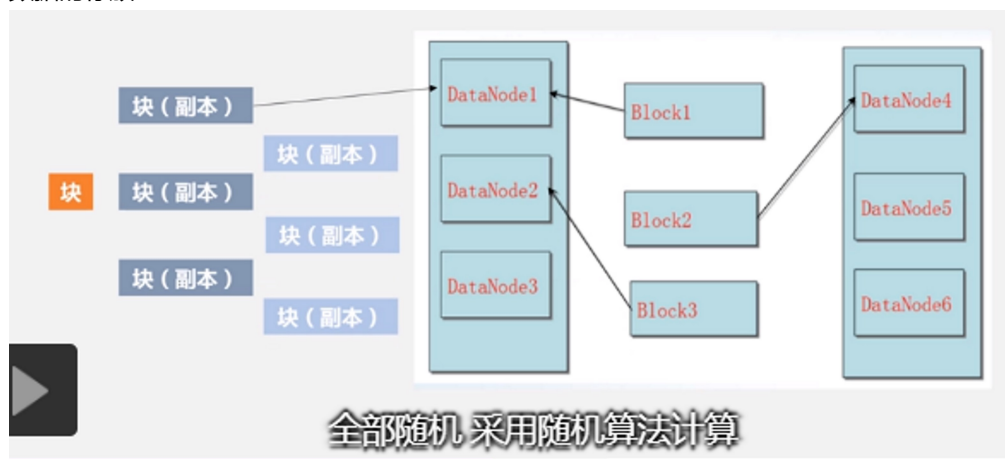
3.4HDFS存储原理

底层因为是廉价机器，所以缺陷是容易不断的出故障，所以有了冗余存储

冗余存储器好处：

- 1.加快数据传输速度
- 2.很容易检查数据错误：因为有很多个副本参照
- 3.保证数据可靠性

数据的存放：



数据的读取：

就近读取，HDFS提供了一个API可以确定一个数据节点所属的机架ID，客户端也可以调用API获取自己所属的机架ID

数据的错误与恢复：

名称节点出错：(FsImage、EditLog) 冷备份从第二名称节点 (SecondaryNameNode) 中恢复，需要暂停一段时间。

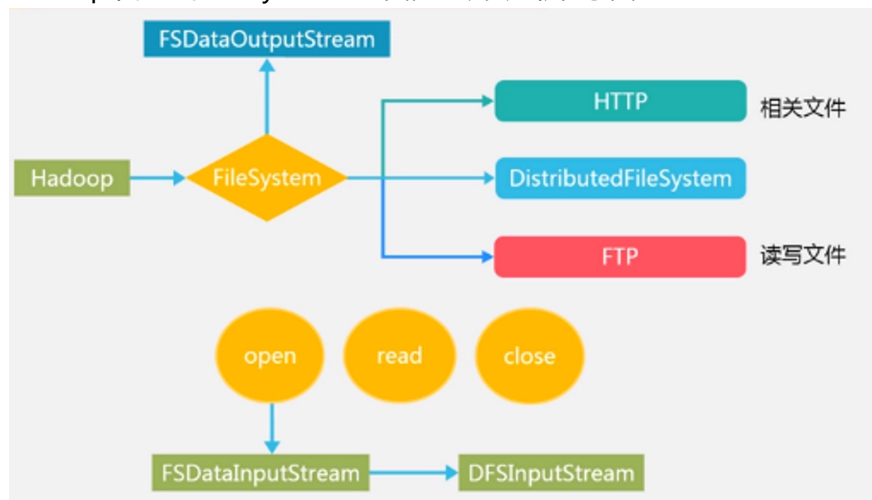
数据节点出错：不断通过远程调用报告给名称节点证明自己还可用，如果没有调用了就标记为宕机，名称节点把备份再备份一份。

数据本身出错：有校验码机制

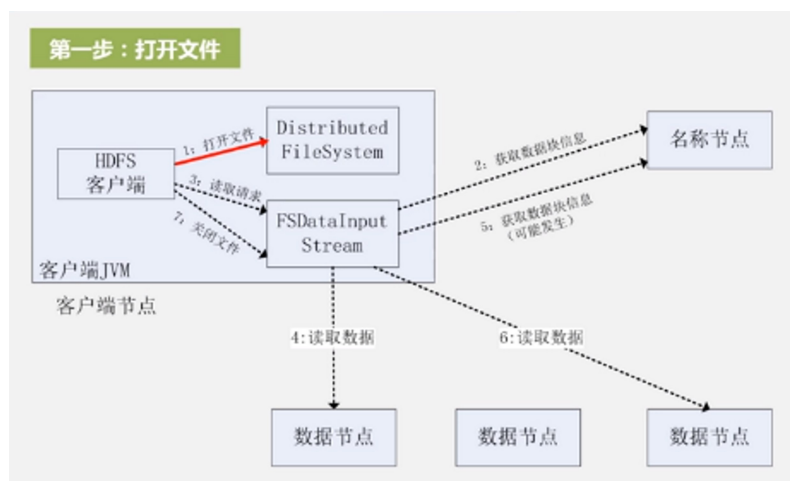
3.5 HDFS数据读写过程

3.5.1 HDFS读数据过程

Hadoop中有一个FileSystem的基类，可以实现很多子类

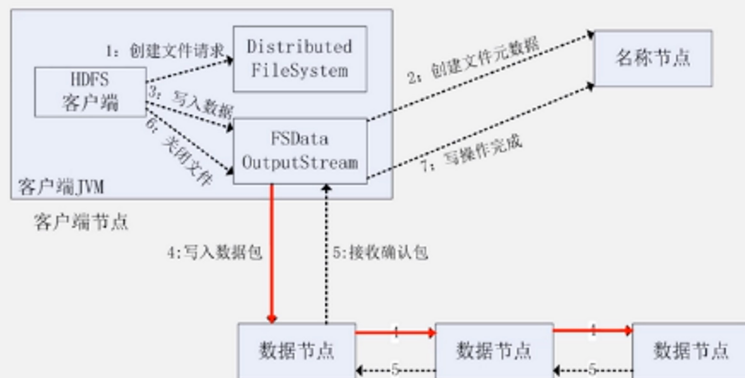


加载hdfs-site.xml、core-site.xml 然后——>读取fs.defaultFS——>hdfs://localhost:9000



3.5.2 HDFS写数据过程

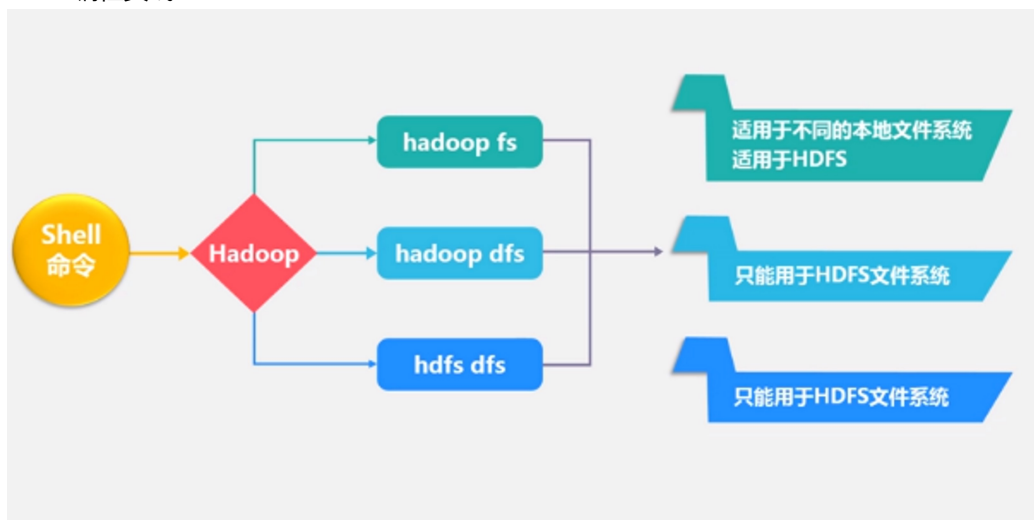
4.写入数据包



沿着一个个节点把它复制下去

- 1.创建文件请求
- 2.创建文件元数据
- 3.写入数据（流水线复制方式 分包）
- 4.写入数据包
- 5.接收确认包
- 6.关闭文件
- 7.写操作完成

HDFS编程实践



`hadoop fs:`

`ls`: 显示置顶文件的详细信息

`mkdir`: 创建相关文件夹

`Cat`: 指定文件内容输出到标准输出

本地文件复制到分布式文件系统HDFS中

`Hadoop fs-cp+本地文件+hdfs`

可通过web界面查看hdfs : <http://localhost:50070> , 一般用shell命令

利用java api与hdfs进行交互

编程实例:

检测伪分布式文件系统HDFS上到底存不存在一个input.txt文件？

第一步：把配置文件放到当前Java工程目录下

需要把集群上的core-site.xml和hdfs-site.xml（这两文件存在/hadoop/etc/hadoop目录下）放到当前工程项目下，即eclipse工作目录的bin文件夹下面

```
$ Wrong FS: hdfs://localhost:9000/user/hadoop/input/input.txt, expected: file:///
```